

## Построение предметной онтологии цифрового пространства научных знаний

*Описан подход к построению предметных онтологий пространства знаний на примере ряда научных дисциплин. В качестве входов в систему предлагается использовать тезаурус ключевых терминов, которые представлены в метаданных информационных ресурсов. Построена структура классификаций и тезаурусов, развитие которой до степени полноценной онтологии предполагает разработку тезауруса смысловых связей тематических рубрик, отражающего онтологические свойства объектов информационного поля.*

**Ключевые слова:** онтология информационного пространства, классификационные системы, тезаурус тематических классов, сопоставление классификаций, тезаурус семантических связей

DOI: 10.36535/0548-0019-2020-12-2

### ВВЕДЕНИЕ

Эффективное использование научно-технической информации предполагает возможность её поиска в пространстве научных публикаций, где она заключена в текстах и подлежит выявлению путём анализа документов. Перспективы автоматического анализа смысла полного текста остаются в настоящее время проблематичными. Альтернативным путём доступа к документальной информации является использование метаданных документа, которые содержат результат интеллектуального анализа его содержания. Метаданные – это ключевые слова и классификационные индексы, которыми снабжают практически все документы в информационном пространстве науки и техники. Необходимость искать сведения среди разобщённых источников ставит задачу согласования средств, с помощью которых систематизированы данные в разнородных ресурсах. Эта задача входит в число актуальных работ по созданию «семантического веба» [1]. Современные поисковые машины ориентированы на поиск по свободной лексике и игнорируют вложенный в создание информационных ресурсов труд индексирования классификационными индексами и ключевыми словами. Из-за недостаточной точности автоматического извлечения смысла из текста в настоящее время невозможна реализация «семантического веба» без использования представления содержания документов средствами классифи-

каций и ключевых слов. Эта задача может быть решена путем формирования сети связей рубрик используемых классификационных систем при привязке к ним ключевых слов, индексирующих документы данной рубрики. Работы в этом направлении начались достаточно давно. В 2008 г. для работы с классификационными системами в ВИНТИ РАН создана база данных [2], в которой содержатся основные классификации, применяемые в пространстве российской научной и технической информации. Методика этой работы и первые результаты её практического применения представлены в монографии [3].

Построение Сети классификационных систем (СКС) ВИНТИ РАН было связано с задачей тематической разметки потока документов при обработке в информационном центре. С этой точки зрения множество классификаций можно условно разделить на три группы:

- *базовые* – обязательно используемые при индексировании публикаций – это ГРНТИ, УДК, Рубрикатор ВИНТИ, Рубрикации информационных продуктов ВИНТИ (РИП);
- *дополнительные* – используемые при описании документов отдельных видов: Номенклатура специальностей научных работников (НСНР, ВАК), Международная классификация патентов (МПК);
- *вспомогательные* – национальные и международные классификаторы, факультативно исполь-

зуемые при определении тематики публикаций: Mathematics Subject Classification (MSC – классификация математических дисциплин Американского математического общества), классификатор журнала Library and Information Science Abstracts (LISA), Общесоюзный классификатор видов экономической деятельности (ОКВЭД), Классификатор по радиационной экологии (КРЭ), Критические технологии Российской Федерации (КТРФ), Приоритетные направления фундаментальных исследований (ПНФИ).

Одновременно с разработкой СКС было начато (совместно с Институтом физики полупроводников – ИФП СО РАН и Библиотекой по естественным наукам – БЕН РАН) подробное исследование смысловых связей рубрик классификационных систем в тематической области физики и электроники [4]. Исследование впервые охватило взаимодействие всего комплекса «обязательных» для отечественной информатики классификаторов – Библиотечно-библиографической классификации (ББК), Универсальной десятичной классификации (УДК) и Государственного рубрикатора научно-технической информации (ГРНТИ).

## ТЕЗАУРУСНАЯ МОДЕЛЬ ОНТОЛОГИИ

Сопоставление нескольких классификаций в области физики полупроводников привело к идее создания тезауруса классификационных рубрик, в котором дескрипторами были бы сами рубрики, а также термины из наименований рубрик разных классификаций [5], действующих в российском информационном пространстве.

Такой тезаурус тематических рубрик по физике полупроводников (ТТРФПП), содержащий около 1600 дескрипторов, был разработан, выложен на сайте научной библиотеки Института физики полупроводников (ИФП) СО РАН (<http://lib.isp.nsc.ru/library/links.php>) и депонирован в ВИНТИ [6]. В тезаурусе каждому дескриптору приписаны коды пяти классификаций: УДК, ББК, ГРНТИ, Рубрикатора ВИНТИ и PACS\* и установлены обычные тезаурусные связи дескрипторов. В дальнейшем тематика физики полупроводников была дополнена терминами смежных областей физики, электроники и нанотехнологий, а модель тезауруса была предложена для построения онтологий других предметных областей [7]. В рамках совместных работ ВИНТИ РАН и БЕН РАН эта модель была реализована как широкая система классификаций и ключевых терминов (КТ) для различных областей знания [8]. В этой системе характер конечного продукта, готового для практического применения, имеют тезаурусы КТ, соответствующие следующим разделам ГРНТИ: *06 Экономика. Экономические науки, 16 Языкознание, 20 Информатика, 29 Физика, 47 Электроника. Радиотехника, 84 Стандартизация*. Эти предметные области были выбраны как экспериментальная база для развития предложенных методов в сферах гуманитарных, информа-

ционных, естественных, технических наук и организационной деятельности.

Тезаурусная модель онтологии информационного пространства по науке и технике была реализована как система «Термин» <http://class.labs.benran.ru> [9–10], в которой представлена семантическая сеть смысловых связей терминов, реально использованных для индексирования содержания научных публикаций. Постановка задачи и начальный этап работ по созданию этой системы описаны в [11]. Отличие такого подхода состоит в том, что если в ТТРФПП термины отобраны исключительно из классификационных таблиц, то в основе лексики БД «Термин» лежат ключевые слова действующих библиографических баз данных – термины, выделенные индексаторами как наиболее важные для описания содержания документов по данной области знания. Такой выбор лексики обеспечивает привязку системы к реальным признакам описания тематики документов.

Для базы данных «Термин» единицами являются понятия, представленные ключевыми словами, связанными с той или иной рубрикой ГРНТИ, ББК или УДК. Понятиям, как правило, даны определения, поясняющие значение термина, которые заимствованы из авторитетных источников, указанных в библиографических ссылках. Между понятиями заданы смысловые связи, установленные полуавтоматическим методом. От термина *A* автоматически даётся «дефинитивная» ссылка на термин *B*, если в дефиниции термина *A* встречается слово *B*, и наоборот. Эти ссылки подлежат экспертному контролю, в ходе которого эксперт указывает тип семантической связи по категориям: совпадение понятий, пересечение понятий, вхождение одного понятия в другое либо отсутствие значимой смысловой связи. При этом около четверти автоматически установленных связей попадает в четвертую категорию и исключается из дальнейшего рассмотрения как установленные на основе случайного совпадения слов, имеющих в данном контексте различный смысл. Например, некоторые термины были связаны с понятием «время» из-за того, что в их определениях встретилось выражение «в то время как».

Экспертное редактирование дефинитивных связей в различных тематических разделах базы данных «Термин» проведено нами на разной глубине анализа. В словарях «Экономика» и «Стандартизация» выполнено наименее трудозатратное редактирование, которое ограничилось только ликвидацией ложных соответствий. Общее количество ложных соответствий во всех тематических разделах базы данных составило около четверти числа связей, предложенных автоматической процедурой. Связями этих словарей можно пользоваться при ориентировочной навигации по источникам данных для повышения полноты информации.

В словарях «Языкознание» и «Информатика» автоматически установленные связи терминов были квалифицированы экспертами по четырём видам логической зависимости, типичным для информационно-поисковых тезаурусов. Для пары терминов может быть указан один из следующих типов связи: 1) эквивалентные – индексируют источники, содержащие данные о том же понятии, что и исходный термин; 2) вышестоящие – индексируют более широкий круг источников; 3) нижестоящие – индексируют часть

\* PACS (Physics and Astronomy Classification Scheme) – классификация знаний по физике и астрономии, разработанная Американским институтом физики (AIP). – URL: <http://www.aip.org/pacs/>.

исходного круга источников; 4) ассоциативные = индексируют источники, круг которых частично пересекается с источниками по исходному термину. Связями терминов в этих словарях можно пользоваться для управления поиском с целью получения либо информации частного характера, либо обобщающих данных, либо сведений о каких-либо сопутствующих явлениях.

При обработке связей терминов в словарях «Физика» и «Электроника» нами решалась задача согласовать связи с онтологической природой понятий с тем, чтобы связями можно было пользоваться для задач фактографического поиска и логического вывода. Для этого отношения между терминами были переосмыслены как отражающие не цели информационного поиска, а явления онтологической реальности.

Наряду со связями терминов друг с другом, во всех словарях установлены связи их с рубриками ГРНТИ, УДК и ББК, которые дополняют систему связей сетью иерархических отношений своих разделов и классов, а также сопоставительными отношениями рубрик разных классификационных систем.

### СОПОСТАВИТЕЛЬНЫЕ ТАБЛИЦЫ КЛАССИФИКАЦИИ

Для разработки предметной онтологии информационного пространства одной из главных является задача тематической навигации среди ресурсов, систематизированных разными классификациями. Для этого сопоставлять классификации удобнее не в тезаурусном, а в табличном формате, когда таблица показывает непосредственно связи рубрик одного классификатора с рубриками других классификаторов. В работе [12] описана методика создания сети классификационных систем ВИНТИ путём полуавтоматического установления прямых связей между рубриками всех классификаций, для которых в базе данных ВИНТИ указано соответствие кодам рубрик ГРНТИ. Методика предусматривает ручное редактирование и пополнение автоматически установленных связей. В экспериментальном режиме такая работа нами выполнена для четырёх разделов ГРНТИ – *16 Языкознание, 20 Информатика, 29 Физика и 47 Электроника. Радиотехника*. Для наиболее важных в библиотечной практике классификаций – ББК, УДК и ГРНТИ – полученные таблицы сопоставлений по некоторым разделам ГРНТИ депонированы в ВИНТИ [13, 14].

Принципы установления соответствий между рубриками различных классификаций соответствуют рекомендациям международного стандарта ISO 25964-2 [15]. Характерная особенность сопоставительной таблицы состоит в избытке соответствий между комбинированными индексами. Это свидетельствует о значительном расхождении классификационных решений в сопоставляемых классификациях. Такие индексы присваивают документам в ходе систематизации для обозначения их тематики, если она захватывает значения различных классов из эталонных таблиц. Когда одна и та же рубрика ГРНТИ описывается комбинированными индексами и в УДК, и в ББК, мы получаем соответствие двух комбинированных классов. Соответствие между простым классом одной классификации и комбинированным классом другой является прямой иллюстрацией расхождения в методах классифицирования.

### ОНТОЛОГИЧЕСКАЯ ИНТЕРПРЕТАЦИЯ ОТНОШЕНИЙ ТЕРМИНОВ

Обращаясь к тезаурусному представлению соответствий, следует отметить, что, несмотря на то что идея тезаурусного представления онтологии отраслей знания восходит к работам над тезаурусом ИФП СО РАН, установление смысловых связей между терминами разделов физики и электроники в системе «Термин» не предполагает автоматического объединения с ТТРФПП. Если тезаурус был ориентирован только на поиск документов, то теперь мы, следуя тенденциям современной информатики, ставим задачу построения в перспективе универсальной онтологии предметной области, которая будет пригодна для решения задач фактографического поиска, управления, извлечения знаний и других задач искусственно-интеллекта. Поэтому в работах над тематикой физики и электроники мы переосмыслили значения применяемых тезаурусных отношений. Они понимаются не в поисковом, а в онтологическом смысле – не как пересечения классов релевантных документов, а как пересечения классов описываемых в документах реалий. Поэтому отношение тождества терминов означает именно тождество денотатов, а не тождество массивов, описывающих их документов. Отношение выше/ниже рассматривается как вхождение класса денотатов одного термина в класс денотатов другого. Это – более строгая интерпретация отношений, чем принято для поиска документов. Если между двумя понятиями *A* и *B* установлены онтологические отношения «*A* эквивалентно *B*» или «*A* включает *B*», то эти отношения будут справедливы и в поисковом смысле, но не наоборот.

Различие между двумя интерпретациями отношений можно пояснить следующим примером. В поисковом смысле антонимы часто можно рассматривать как эквиваленты; например, все статьи о неустойчивости плазмы релевантны запросу об устойчивости плазмы, поскольку критерии этих явлений совпадают, и их определения равно соответствуют и тому, и другому понятию. Но в онтологическом смысле эти понятия исключают друг друга и не могут рассматриваться как эквивалентные или пересекающиеся.

Что же касается «ассоциативного» отношения *A*×*B*, то в онтологическом смысле оно понимается как наличие у денотатов общих атрибутов или как смежность соответствующих реалий в пространстве и времени, что вполне сходится с пониманием этого отношения в поисковом смысле как пересечения массивов релевантных документов.

Базовый массив связей терминов в системе «Термин», в первом приближении, был получен автоматически на основе «дефинитивного» поиска соответствий (употребление одного термина в составе дефиниции другого). Такой поиск не даёт возможности квалифицировать найденную связь по категориям видов тезаурусных отношений. Уточнение вида автоматически установленной связи по категориям «совпадение – вхождение – пересечение» осуществлялось методом интеллектуального анализа. При этом существенную долю дефинитивных связей (примерно четверть) пришлось исключить как установленные из-за формального совпадения слов с семантически не связанными значениями.

## Онтологические связи терминов в словарях «Физика» и «Электроника»

Знак	Чтение	Значение
$A = B$	совпадает, равно, тождественно	Термины А и Б обозначают тождественные или почти тождественные множества реалий (синонимы).
$A \gg B$	больше, шире, включает	Термин А обозначает множество реалий, в которое включено множество реалий, обозначаемых термином Б, причём объёмы этих множеств соизмеримы.
$A \ll B$	меньше, уже, входит в	Термин А обозначает множество реалий, включённое во множество реалий, обозначаемых термином Б, причём объёмы этих множеств соизмеримы.
$A \times B$	пересекается с	Множества реалий, обозначаемые терминами А и Б, пересекаются в существенной части.
$A - B$	дефинитивно связан с	Реалии, обозначаемые терминами А и Б, связаны прагматическими связями, но их множества, вероятно, не пересекаются, относясь к различным онтологическим категориям.

Но и среди сохраненных действительных связей не все целесообразно применять при обычном поиске документов. Например, понятия «автомобиль» и «бензин» явно связаны, и это может быть отражено в их определениях. Но использовать документы о бензине как релевантные для поиска документов об автомобилях (и наоборот) вряд ли целесообразно в общем случае. Однако, если в поисковой системе будут реализованы специфические модальности поиска «источник энергии для» или «применяется в», то связь таких терминов будет востребована. Поэтому такие связи не ликвидированы, а оставлены как особая категория «слабых пересечений» в качестве кандидатов на установление специфических режимов поиска, учитывающих сущностные отношения объектов онтологической реальности. В табл. 1 показаны виды связей терминов, установленные в словарях «Физика» и «Электроника» системы «Термин».

В настоящее время нами разрабатывается система свойств объектов, которая позволит дифференцировать «дефинитивные» связи на основе типа отношений между определяемым понятием и понятиями в определениях. В первую очередь, предполагается установить следующие связи понятий, типичные для многих известных онтологий верхнего уровня:

элемент / множество (например, *Земля / планета*);  
 компонент / система (например, *атомное ядро / атом*);  
 атрибут / носитель атрибута (например, *текучесть / жидкость*);  
 процесс / результат (например, *конденсация / конденсат*);  
 действие / действующее начало (например, *испарение / испаритель*);  
 объект / локализация (например, *орбитальный электрон / электронная орбита*).

Специфические свойства объектов информации в сфере физики и электроники должны отражать такие связи как:

измеритель / измеряемая характеристика;

генератор / генерируемая сущность;  
 фундаментальное взаимодействие / элементарная частица;  
 вещество / физические свойства;  
 излучение / волновые свойства;  
 тело / геометрические характеристики.

В других тематических доменах следует установить свои специфические отношения, определяемые онтологическими свойствами реалий домена. Так, для словаря «Языкознание» может быть использована система связей дескрипторов, зафиксированная в тезаурусе С. Е. Никитиной [16].

Отношения, отражающие взаимодействие объектов реальности, в свою очередь, образуют иерархическую структуру наподобие структуры тезауруса названий объектов. В совокупности тезаурус объектов и тезаурус отношений будут представлять онтологию предметной области в аспекте её состава (номенклатуры) и в аспекте взаимоотношений объектов в процессах физических явлений.

## ЗАКЛЮЧЕНИЕ

Результатом работы коллектива специалистов ВИНТИ РАН является онтология пространства научной информации, представленная двумя сопряженными семантическими сетями на одном поле физических реалий, – тезаурусами ключевых терминов 63 предметных областей и таблицами соответствий основных классификационных систем. Доступ к навигации по этим сетям осуществляется на сайте: <http://rffi.viniti.ru>.

Состав предметных областей онтологии определяется верхним уровнем ГРНТИ. Параметры словарей (тезаурусов), сформированных для этих областей, представлены в табл. 2. Здесь в первой колонке приводится формулировка рубрики ГРНТИ, во второй колонке – количество включённых в словарь терминов, характеристика установленных смысловых связей терминов даётся по следующим катего-

риям: «дефинитивные» – установленные автоматической процедурой без существенного экспертного редактирования в большинстве тематических областей; «ассоциативные» – связи, прошедшие экспертный контроль на соответствие задачам смысловой навигации по областям «Экономика» и «Стандартизация»; «тезаурусные» – связи, квалифицированные экспертами по трем видам обычных тезаурусных от-

ношений терминов областей «Языкознание» и «Информатика»; «онтологические» – связи, иерархия которых отражает онтологические категории реалий действительности в областях «Физика» и «Электроника». Всего в «Термин» введено более 12 тыс. терминов, которые связаны примерно 300 тысячами отношений друг с другом и приписаны к рубрикам трёх классификаций – ГРНТИ, УДК и ББК.

Таблица 2

Тезаурусы системы «Термин» по разделам ГРНТИ

Рубрика ГРНТИ	Число терминов
• <u>02 Философия</u>	92
• <u>03 История и исторические науки</u>	79
• <u>04 Социология</u>	120
• <u>05 Демография</u>	74
• <u>06 Экономика и экономические науки</u>	202
• <u>10 Государство и право. Юридические науки</u>	149
• <u>11 Политика и политические науки</u>	176
• <u>12 Науковедение</u>	98
• <u>13 Культура. Культурология</u>	96
• <u>14 Народное образование. Педагогика</u>	102
• <u>15 Психология</u>	140
• <u>16 Языкознание</u>	531
• <u>17 Литература. Литературоведение. Устное народное творчество</u>	54
• <u>18 Искусство. Искусствоведение</u>	67
• <u>19 Массовая коммуникация. Журналистика. Средства массовой информации</u>	48
• <u>20 Информатика</u>	392
• <u>21 Религиоведение</u>	54
• <u>27 Математика</u>	269
• <u>28 Кибернетика</u>	75
• <u>29 Физика</u>	719
• <u>30 Механика</u>	81
• <u>31 Химия</u>	198
• <u>34 Биология</u>	714
• <u>36 Геодезия. Картография</u>	110
• <u>37 Геофизика</u>	26
• <u>38 Геология</u>	198
• <u>39 География</u>	105
• <u>41 Астрономия</u>	169
• <u>44 Энергетика</u>	146
• <u>45 Электротехника</u>	127
• <u>47 Электроника. Радиотехника</u>	488
• <u>49 Связь</u>	209
• <u>50 Автоматика. Вычислительная техника</u>	170
• <u>52 Горное дело</u>	101
• <u>53 Металлургия</u>	251
• <u>55 Машиностроение</u>	761
• <u>58 Ядерная техника</u>	112
• <u>59 Приборостроение</u>	114
• <u>60 Полиграфия. Репрография. Фотокинетехника</u>	117
• <u>61 Химическая технология. Химическая промышленность</u>	372
• <u>62 Биотехнология</u>	212
• <u>64 Легкая промышленность</u>	216
• <u>65 Пищевая промышленность</u>	302
• <u>66 Лесная и деревообрабатывающая промышленность</u>	183
• <u>67 Строительство. Архитектура</u>	161
• <u>68 Сельское и лесное хозяйство</u>	90

Рубрика ГРНТИ	Число терминов
• 69 Рыбное хозяйство. Аквакультура	170
• 70 Водное хозяйство	197
• 71 Внутренняя торговля. Туристско-экскурсионное обслуживание	126
• 72 Внешняя торговля	132
• 73 Транспорт	239
• 75 Жилищно-коммунальное хозяйство. Домоводство. Бытовое обслуживание	92
• 76 Медицина и здравоохранение	134
• 77 Физическая культура и спорт	139
• 78 Военное дело	131
• 82 Организация и управление	134
• 83 Статистика	149
• 84 Стандартизация	109
• 85 Патентное дело. Изобретательство. Рационализаторство	49
• 86 Охрана труда	99
• 87 Охрана окружающей среды. Экология человека	377
• 89 Космические исследования	167
• 90 Метрология	96

Таблица 3

### Состав Сети классификационных систем ВИНТИ РАН

Наименование классификации	Общее количество рубрик (классов)
Универсальная десятичная классификация (УДК)	166878
Библиотечно-библиографическая классификация (ББК) – верхний уровень иерархии	25200
Государственный рубрикатор научно-технической информации (ГРНТИ)	8336
Международная патентная классификация (МПК) – верхний уровень иерархии	778
Рубрикатор Российского научного фонда (РНФ)	603
Рубрикатор Российского фонда фундаментальных исследований (РФФИ)	546
Номенклатура специальностей научных работников высшего уровня (Номенклатура ВАК)	502
Рубрикатор журналов Scopus издательства Elsevier	336
Рубрикатор библиографической базы данных Web of Science (WoS)	265
Рубрикатор «Fields of Science and Technology» Организации экономического содействия и развития	49
Итого	203493

Сеть соответствий тематических рубрик представлена взаимными отражениями друг на друга десяти классификационных систем, перечень и объёмы которых указаны в табл. 3.

Семантическая сеть находится в состоянии актуализации, редактирования, корректировки и развития сотрудниками ВИНТИ РАН. Опыт создания и ведения тематического фрагмента сети по тематике «Электроника» представлен в докладе на конференции LIBWAY [17] и в публикации [18].

Таким образом, в результате работ большого коллектива специалистов, возглавляемого ВИНТИ РАН, для области научной и технической информации в пространстве терминов и тематических рубрик создана семантическая сеть, связывающая входные точки тематического поиска в различных информационных источниках, наделённых разными системами доступа – классификационными и предметными

индексами. Сеть терминов представлена в тезаурусном и в табличном вариантах, каждый из которых объединяет смысловыми связями элементы разнородных классификационных систем. Созданная семантическая сеть может стать основой для построения среды информационных ресурсов Единого цифрового пространства научных знаний [20].

На текущем этапе работ для внедрения предлагаемой семантической сети в промышленную эксплуатацию необходимо:

- объединить базы данных «Термин» и Сеть классификационных систем на единой программно-технологической основе;
- осуществить каталогизацию и классификацию типов ассоциативных и дефинитивных связей терминов и разработку на этой основе тезауруса отношений терминов, которые задают аксиомы для объектов онтологии в каждом тематическом домене;

- ввести в базу данных дополнительно полные таблицы всех рассматриваемых классификаций;
- пополнить сеть ключевыми словами наиболее важных (наиболее цитируемых) работ;
- обеспечить автоматическое управление связями сети при внесении изменений в состав объектов онтологии.

\* \* \*

Авторы выражают благодарность всем участникам работ по созданию Сети классификационных систем и терминологических тезаурусов «Терм» в базах данных ВИНТИ РАН и БЕН РАН.

## СПИСОК ЛИТЕРАТУРЫ

1. Bast H., Buchhold B., Haussmann E. Semantic search on text and knowledge bases // *Fundations and Trends in Information Retrieval*. – 2016. – 10 (2-3). – С. 119-271. – DOI: 10.1561/15000000032.
2. Шапкин А.В. Практические вопросы построения системы классификационных схем // *Научно-техническая информация. Сер. 2*. – 2006. – № 6. – С. 1-14.
3. Гиляревский Р.С., Шапкин А.В., Белоозеров В.Н. Рубрикатор как инструмент информационной навигации. – Санкт-Петербург : Профессия, 2008. – 352 с.
4. Белоозеров В.Н., Шабурова Н.Н. Классификационные системы как средство поиска информации по физике полупроводников // *Библиосфера*. – 2008. – № 3. – С.34–42.
5. Белоозеров В.Н., Шабурова Н.Н. Сопоставительный тезаурус классификационных систем по физике полупроводников // *Информационное обеспечение науки: новые технологии : Сборник научных трудов*. – Москва : Научный мир, 2009. – С. 311–322.
6. Белоозеров В.Н., Шабурова Н.Н. Тезаурус тематических рубрик по физике полупроводников // *Депонировано в ВИНТИ 2013-12-24, № 379-B2013*. – Москва, 2013.
7. Белоозеров В.Н., Шабурова Н.Н. Тезаурус библиографических классификаций как модель интеграции информационных ресурсов // *Международная конференция «Информационное общество: Состояние и тенденции межгосударственного обмена научно-технической информацией в СНГ»*, Москва, 27-28 окт. 2011. – Москва : ВИНТИ, 2011. – С. 18–22.
8. Сютюренко О.В. [и др.] Сеть классификаций по науке и технике как механизм смысловой навигации и поиска информации в пространстве знаний // *Депонировано в ВИНТИ РАН 19.12.2019, № 120-B2019*. – Москва, 2019. – 40 с.
9. Якшин М.М., Калёнов Н.Е. Классификаторы: создание базы данных терминологических словарей // *Сб. научных трудов «Информационное обеспечение науки: новые технологии»*. – Москва : БЕН РАН, 2015. – С. 137-146.
10. Калёнов Н.Е., Сенько А.М. Программная оболочка системы терминологических словарей // *Научное издание международного уровня – 2019: стратегия и тактика управления и развития : Материалы 8-й Международной научно-практической конференции*, г. Москва, 23-26 апреля 2019 г. – Екатеринбург : Изд-во Уральского университета, 2019. – С. 58-63.
11. Kalenov N.E., Senko A.M. Interactive system of terminological dictionaries as one of the elements in the ontology of scientific knowledge // *Software Journal: Theory and Applications (electronic Journal)*, – 2019. – Iss. 4. DOI: 10.15827/2311-6749.33.423
12. Антопольский А.Б. [и др]. Разработка онтологии информационного пространства знаний на основе дефинитивных связей // *Научно-техническая информация. Серия 1*. – 2017. – № 11. – С. 19–24; Antopol'skii A.B. [et al.]. The development of a semantic network of keywords based on definitive relationships // *Scientific and Technical Information Processing*. – 2017. – 44 (4). – 261–265.
13. Антошкова О.А. [и др]. Методика построения онтологии информационных ресурсов в виде сети библиографических классификаций // *Научно-техническая информация. Сер. 1*. – 2017. – № 11. – С. 25–30; Antoshkova O. A. [et al.]. On a method for constructing an ontology of scientific and technical information as a network of bibliographic classifications // *Scientific and Technical Information Processing*. – 2017. – № 44(4). – P. 266–275.
14. Белоозеров В.Н., Шабурова Н.Н., Смирнова О.В. Сопоставительные таблицы классификационных систем УДК и ББК по тематике «Радиоэлектроника и физика твёрдых тел» // *Депонировано ВИНТИ РАН 14.01.18 № 9-B2018*. – Москва, 2018. – 38 с.
15. Белоозеров В.Н., Маркарова Т.С., Смирнова О.В., Астахова Т.С. Сопоставительные таблицы классификационных систем УДК, ББК и ГРНТИ по тематике «Языкознание. Языки» // *Депонировано ВИНТИ РАН 24.01.18 № 10-B2018*. – Москва, 2018. – 44 с.
16. ISO 25964-2:2013 Information and documentation - Thesauri and interoperability with other vocabularies – Part 2: Interoperability with other vocabularies : International standard / ISO. – Geneva, Switzerland, 2013. – 99 p.
17. Никитина С.Е. Тезаурус по теоретической и прикладной лингвистике (Автоматический анализ текста). – Москва : Наука, 1978. – 375 с.
18. Шабурова Н.Н. Опыт работы с БД «TERMIN» // *Международная научно-практическая конференция «Наука, технологии и информация в библиотеках» (LIBWAY–2018)*, Новосибирск, 12-15 сентября 2018 г. – URL : <https://www.libway.ru/program/Saeffd4c7045cd0be895b01a>.
19. Shaburova N.N. Operational experience in DB “TERMIN” // *Journal of Information Science Theory and Practice*. – 2019. – Vol. 7(3). – P. 21-30. DOI : <https://doi.org/10.1633/JISTaP.20197.3.2>.

20. Антопольский А.Б. [и др.] Принципы построения и структура Единого цифрового пространства научных знаний (ЕЦПНЗ) // Научно-техническая информация. Сер. 1. – 2020. – № 4. – С. 9–17.

*Материал поступил в редакцию 28.05.20.*

#### **Сведения об авторах**

**БЕЛООЗЕРОВ Виктор Николаевич** – кандидат филологических наук, доцент, ведущий научный сотрудник ВИНТИ РАН

**ДМИТРИЕВА Елена Юрьевна** – кандидат технических наук, заведующая отделением ВИНТИ РАН  
e-mail niipio@viniti.ru

**КАЛЕНОВ Николай Евгеньевич** – доктор технических наук, профессор, главный научный сотрудник Межведомственного суперкомпьютерного центра РАН  
e-mail nekalenov@mail.ru

**ШАБУРОВА Наталья Николаевна** – кандидат педагогических наук, заведующая научной библиотекой Института физики полупроводников СО РАН  
e-mail shaburova@isp.nsc.ru

**ШАПКИН Александр Владимирович** – кандидат технических наук, заведующий отделом ВИНТИ РАН  
e-mail ss@viniti.ru